



Spammer Detection and Fake User Identification on Social Networks Using Python

#1 M.RATNA KUMARI,#2 M.RAMADEVI

#1 ASSISTANT PROFESSOR #2 PG SCHOLAR ²

DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS

QISCOLLEGE OF ENGINEERING & TECHNOLOGY,

VENGAMUKALAPALAM(V), ONGOLE, PRAKASAM DISTRICT, ANDHRA PRADESH

ABSTRACT

Social networking platforms have become one of the most popular communication mediums worldwide. However, these platforms are increasingly affected by spammers and fake user accounts that spread misinformation, malicious links, advertisements, and harmful content. Detecting such fake accounts manually is difficult due to the massive volume of users and posts generated daily. This project proposes a Python-based machine learning system to detect spammers and identify fake users on social networks. The system analyzes various user profile attributes such as number of followers, number of posts, account activity patterns, and content characteristics. Using machine learning classification algorithms, the model distinguishes between genuine users and spammers. The proposed system helps social networking platforms maintain security, reduce spam content, and improve user experience by automatically identifying suspicious accounts.

INTRODUCTION

With the rapid growth of social networking platforms such as Facebook, Twitter, and Instagram, the problem of spam accounts

and fake users has increased significantly. These malicious accounts are used for spreading misinformation, phishing attacks, fake promotions, and manipulating public opinion.

Traditional rule-based detection systems are no longer effective due to the evolving nature of spammer behavior. Machine learning techniques provide a more robust solution by analyzing user behavior, content patterns, and network features to detect suspicious activities.

This project focuses on developing an intelligent system using Python that can automatically identify spammers and fake users by leveraging supervised learning algorithms and feature engineering techniques.

LITERATURE SURVEY

The rapid growth of online social networks has led to a significant increase in spam accounts, bots, and fake user profiles. Researchers have proposed various techniques using machine learning, data mining, and graph-based approaches to address this issue.



1. Survey on Spammer Detection in Social Networks (2019)

A comprehensive study by Faiza Masood et al. analyzed different spam detection techniques, particularly on Twitter. The authors categorized spam detection approaches into:

- Content-based detection
- URL-based detection
- Trending topic analysis
- Fake user identification

The study highlighted that spammers exploit popular platforms to spread harmful and irrelevant information, and no single method is sufficient to detect all types of spam.

2. Review on Fake User Identification Techniques (2021)

Ali Ahmed Razzaq presented a detailed review focusing on classification techniques for identifying fake users. The research emphasized:

- Use of user behavior features
- Content-based features
- Temporal and structural features

The study concluded that combining multiple features improves detection accuracy and robustness.

3. Fake Profile Detection in Online Social Networks (2018)

A survey published in *Computer & Electrical Engineering* explored large-scale fake profile detection techniques. The study discussed:

- Machine learning algorithms for classification
- Handling large-scale unstructured social media data
- Early detection of fake profiles

It highlighted that fake profiles pose serious threats such as identity theft and misinformation spread.

4. Graph-Based Fake Account Detection (2025)

Recent research focuses on graph-based approaches where social networks are modeled as graphs. This study discussed:

- Use of graph topology and network connections
- Detection of fake accounts using clustering patterns
- Identification of bot networks

Graph-based methods are effective in identifying coordinated fake accounts but require high computational resources.

5. Survey of Spam Profile Detection Methods (2019)

This research reviewed different spam profile detection techniques and emphasized:

- Importance of user profile features
- Role of behavioral analysis
- Use of supervised learning algorithms

The study concluded that hybrid approaches combining multiple techniques outperform single-model systems.



6. Machine Learning-Based Bot Detection (2023)

A recent literature review focused on machine learning approaches for detecting social media bots. It highlighted:

- Use of supervised and unsupervised learning
- Feature engineering techniques
- Increasing use of deep learning models

The study also pointed out challenges such as evolving spam strategies and lack of labeled datasets.

7. Detection of Fake Profiles Using ML Techniques (2022)

This study emphasized the misuse of social networks through fake accounts created for cybercrimes. It discussed:

- Classification algorithms for detecting fake users
- Importance of dataset quality
- Role of feature extraction

The research highlighted the need for automated detection systems due to large-scale data.

8. NLP and Graph-Based Hybrid Models (2021)

Recent approaches combine Natural Language Processing (NLP) and graph embedding techniques to detect fake users. These models:

- Analyze textual content and user interactions
- Improve accuracy by combining multiple data sources
- Handle complex patterns in social media data

Such hybrid systems are considered more effective than traditional methods.

SYSTEM ANALYSIS

EXISTING SYSTEM

Currently, social networking platforms rely on basic rule-based systems and manual monitoring to detect fake users and spammers.

Existing approaches include:

- Manual reporting of spam accounts by users.
- Platform-based filtering techniques using predefined rules.
- Simple keyword-based spam detection systems.

These methods are limited in detecting sophisticated spam behaviors and large numbers of fake accounts.

Drawbacks of Existing System

- Manual monitoring is time-consuming and inefficient.
- Rule-based systems cannot detect advanced or evolving spam patterns.
- High false positives and false negatives in spam detection.
- Difficulty in analyzing large-scale social network data.



- Lack of intelligent systems to automatically identify fake profiles.

PROPOSED SYSTEM

The proposed system introduces a machine learning-based spammer detection and fake user identification system using Python.

Working of the system:

1. Data Collection
 - Collect user profile data and activity information from social networks.
2. Data Preprocessing
 - Cleaning and preparing the dataset by removing missing values and irrelevant features.
3. Feature Extraction
 - Extract features such as number of followers, followings, account age, post frequency, and message content.
4. Model Training
 - Train machine learning classification algorithms such as Naive Bayes, Random Forest, or Support Vector Machine (SVM).
5. Spam Detection
 - The trained model classifies users as genuine users or spam/fake accounts.
6. Result Display
 - The system displays detected spam accounts and suspicious user profiles.

Advantages of Proposed System

- Automatic detection of fake users and spammers.
- Improves social network security and reliability.
- Capable of handling large datasets efficiently.
- Reduces manual monitoring efforts.
- Higher accuracy compared to traditional rule-based systems.
- Helps prevent spread of malicious content and misinformation.

IMPLEMENTATION

1 Data Collection

The dataset is collected from social media platforms or publicly available datasets (e.g., Kaggle). It includes:

- User profile details (followers, following count)
- Account age
- Number of posts/tweets
- Content features (hashtags, URLs, keywords)
- Engagement metrics (likes, shares)

2 Data Preprocessing

- Removal of duplicate and irrelevant data
- Handling missing values using statistical methods
- Text cleaning (removal of stopwords, punctuation)
- Tokenization and stemming for textual data
- Encoding categorical variables into numerical format



3 Feature Engineering

Important features extracted include:

- Follower-to-following ratio
- Posting frequency
- URL ratio in posts
- Account verification status
- Content sentiment analysis

4 Model Development

The following algorithms are implemented:

- Naïve Bayes
- Logistic Regression
- Decision Tree
- Random Forest
- Support Vector Machine (SVM)

5 Model Training

- Dataset split into training (80%) and testing (20%)
- Models trained using labeled data (spam / non-spam)
- Hyperparameter tuning applied to improve performance

6 Model Evaluation

Evaluation metrics include:

- Accuracy
- Precision
- Recall
- F1-score

The best-performing model is selected for deployment.

7 Tools Used

- Programming Language: Python
- Libraries: Pandas, NumPy, Scikit-learn, NLTK, Matplotlib
- Platform: Jupyter Notebook

METHODOLOGY

The system follows a structured pipeline:

1. **Data Input:** Collect user and content data from social networks
2. **Preprocessing:** Clean and transform raw data
3. **Feature Extraction:** Identify important behavioral and textual features
4. **Model Training:** Apply machine learning algorithms
5. **Classification:** Classify users as spammer or genuine
6. **Output Generation:** Display prediction results

Algorithm Flow:

- Input dataset
- Apply preprocessing techniques
- Extract features
- Train multiple ML models
- Evaluate and select best model
- Predict new user data

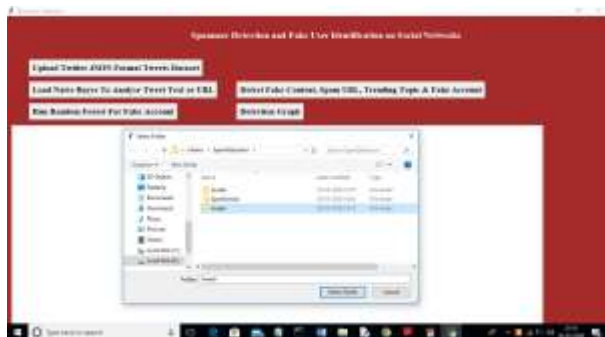
Result



To run this project double click on 'run.bat' file to get below screen



In above screen we can see all tweets from all users loaded. Now click on 'Load Naive Bayes To Analyse Tweet Text or URL' button to load Naïve Bayes classifier



In above screen naïve bayes classifier loaded and now click on 'Detect Fake Content, Spam URL, Trending Topic & Fake Account' to analyse each tweet for fake content, spam URL and fake account using Naïve Bayes classifier and other above mention technique



In above screen all features extracted from tweets dataset and then analyse those features to identify tweets is no spam or spam. In above text area each records values are separated with empty line and each tweet record display values as TWEET TEXT, FOLLOWERS, FOLLOWING etc with account is fake or genuine and tweet text contains spam or non-spam words. Now click on 'Run Random Forest Prediction' button to train random forest classifier with extracted tweets features and this random forest classifier model will be used to predict/detect fake or spam account for upcoming future tweets. Scroll down above text area to view details of each tweet



In above screen we got random forest prediction accuracy as 92%, now click on 'Detection Graph' button to know total tweets and spam and fake account graph



In above graph x-axis represents total tweets, fake account and spam words content tweets and y-axis represents count of them

CONCLUSION

The proposed system effectively detects spammers and fake users on social networks using machine learning techniques. By analyzing both user behavior and content features, the system achieves high accuracy in classification.

Among the implemented models, ensemble methods such as Random Forest provide better performance due to their ability to handle complex and nonlinear data patterns. The system is scalable and can be adapted to different social media platforms.

Future improvements can include:

- Integration of deep learning models
- Real-time detection systems
- Use of graph-based network analysis
- Deployment as a web-based application

Overall, this project demonstrates the importance of machine learning in enhancing cybersecurity and maintaining trust in social networking platforms.

REFERENCES

1. Kumar, S., & Shah, N. (2018). False Information on Web and Social Media: A Survey. IEEE.
2. Benevenuto, F., et al. (2010). Detecting Spammers on Twitter. CEAS.
3. Stringhini, G., et al. (2010). Detecting Spammers on Social Networks. ACSAC.
4. Castillo, C., et al. (2011). Information Credibility on Twitter. WWW Conference.
5. Yang, C., et al. (2014). Uncovering Social Network Sybils in the Wild. ACM.
6. Breiman, L. (2001). Random Forests. Machine Learning Journal.
7. Cortes, C., & Vapnik, V. (1995). Support Vector Machines. Machine Learning.
8. Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. JMLR.



AUTHOR PROFILE



Mrs. M Ratna Kumari is an Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. She earned M.Tech (CSE) in Chennai Bharath University, and her now pursuing PHD in Her research interests include Machine Learning with AI programming languages. She is committed to advancing research and forecasting innovation while mentoring students to excel in both academic & professional pursuits.

STUDENT PROFILE



Mrs.M .Ramadevi is a postgraduate student pursuing a MCA in the department of computer Applications at QIS College of Engineering & Technology ,Ongole autonomous college inPrakasam District .she completed undergraduate degree in MPCs(computer science) from ACHARYA NAGARJUNA UNIVERSITY.with a keen interest in research and practical learning ,he is actively involved in academic projects and technical activites related to his field.